2009

Attitudes Toward Automation and Information Requirements of Experienced Predator Operators

William P. Marshak

John Hudson

Jeff Collier

Attitudes Toward Automation and Information Requirements of Experienced Predator Operators

William P. Marshak, John Hudson, Jeff Collier
SYTRONICS, Inc.
Michael C. Findler
SYTRONICS, Inc. and Wright State University

Two simulation studies were conducted using the Multiple Uninhabited Air Vehicle Agency (**MAGE**) software with nine Predator experienced pilots and sensor operators. In one study, MAGE's situation awareness and decision support elements were employed as supplemental information displays to a Predator crew station simulation within the Air Force's SIMAF facility. In the second simulation, each pilot controlled two autonomous UAVs during a complex multi-target mission scenario using the MAGE software suite. After the simulations were complete, the crewpersons answered questions about their attitudes concerning automated features and what kinds of non-vehicle information (weather, air spaces, intelligence, etc.) they felt they needed during operational missions. The results reflected their experience with the Predator system, to some degree their age, and the particulars of the MAGE user interface. Their responses provided valuable insights about the information requirements and evolution of the UAV Ground Control Station's user interface.

In response to operational needs observed in the Predator operations community, ***Air Force Research Laboratory (AFRL)*** identified two specific areas of deficiency and directed that research to address the needs. The first need was to enhance ***Uninhabited Air System (UAS)*** performance by integrating net-centric information from beyond the ***Uninhabited Air Vehicle (UAV)*** into its ***Ground Control System (GCS)***. Operations units had already added additional LCD displays to the Predator GCS to provide supplemental information into the GCS. The second operational need was to reduce Predator manning requirements by allowing a pilot to control more than one UAV at a time. MAGE was created to study how to accomplish these needs (Figure 1).

Figure 1. MAGE Hardware Configuration

"Operational" Test and Evaluation of MAGE as a Supplemental Display in SimAF

The installation and evaluation of MAGE as a supplemental display in the SimAF MQ-9 Simulator was performed for two reasons: (1) to evaluate technical integration issues and (2) to evaluate human system interface issues. Only the human system evaluation will be discussed here. Transitioned MAGE technology would most likely be as supplemental displays to the UAV manufacturer's GCS. The integration of MAGE into the SimAF MQ-9 Simulator would provide valuable insights into the integration strategies and problems associated with such integration. Since the opportunity to integrate into an operational GCS did not present itself, the integration into an existing high fidelity simulation was the next best thing.

The second reason for the integration had to deal with evaluating how operators would employ MAGE supplemental displays. The three ***Subject Matter Expert (SME)*** operators were given training in the MAGE displays and employed the MAGE displays during the SimAF simulation and a series of stand-alone target engagements during lulls in the SimAF simulation. Observers noted the operations and comments during the exercise to provide insights into both the MAGE capabilities and their use during mission simulation. One of the SME's pointed out there had not been any systematic requirements analysis for supplemental displays in the Predator GCS that he was aware of.

At the conclusion of the day, an exit interview was taken from each of the three crew persons concerning their experience with the MAGE displays in the Reaper crew station simulation. Question 1 simply asked whether supplemental information was of value to Predator crews. The answer was a unanimous yes among the SMEs. Questions 2 through 11 asked the three SMEs to rate the importance of different information sources and is shown in Figure 2.

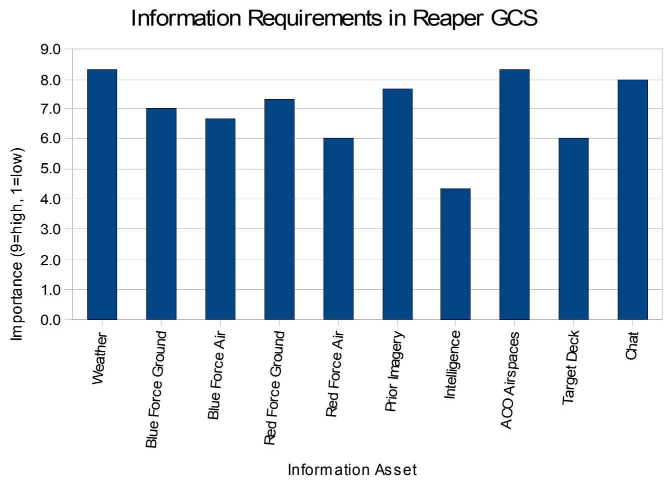


Figure 2. Average rated importance of information sources in Reaper/Predator operations (N=3)

The two highest rated categories (>8) were weather and Air Operations Center airspaces, which relate to the safe operation of the UAV. The next most important categories (7-8) included red force ground tracks, prior imagery, and chat. These are all related to prosecution of ground targets. The next highest categories (6-7) include other tracks (air and ground), and the overall target deck. Intelligence rated the lowest 4.3. This is due to the fact that target assignment is not the crews job; they only have veto over execution and then only if the rules of engagement or commander's intent would be violated by the strike.

All three SMEs agreed that there is a very real need for supplemental information in the GCS. Only one suggested an additional source of information, which was fuel states and range information when dealing with requests to retask the vehicle they were flying. This information would require coupling of the tracker fuel information and the FalconView-Mission Planner combination. When asked about database investment (Question 13), two of the three SMEs were aware of efforts to create mission and intelligence databases. Specifically called out were air tracks they shared the airspace with and imagery. When specifically asked about the TD, the SME's liked the augmentations of FalconView with qualifications. They thought the MAGE TD was an improvement over the current augmented FalconView displays, but thought it needed to be tuned more toward the Predator mission specific needs. There was interest in real-time ground track information to maintain situation awareness of friendly forces. There was concern expressed about clutter and clutter management. Clearly, fielding of the augmented TD will require revisiting operational units to further define the kinds of information displays to satisfy their needs.

Voice recognition was deemed useful, but our implementation was somewhat problematic. Even though interviews with one of the three SMEs were used to define the vocabulary, there was enough within user variation not to mention between user variations to create dissatisfaction. Clearly, the brief training opportunity was insufficient to familiarize the SMEs with the specific vocabulary. Broader vocabulary definition and greater recognition accuracy were both requested. A desire for dialog recognition was stated for use in chat operations since keyboard use demands skill and attention. The younger SME, the SO, was much more receptive to voice input and asked that coordinate definition function be added. Since all three operators were knowledgeable in FalconView's manual control, they often resorted to those manual controls instead of voice control.

Subjects were asked specifically about the effectiveness of adding additional special purpose supplemental displays, as is the case in the operational GCSs which various reports now place at six. There was consensus that the current configuration was getting the job done, but at some expense of operator workload. There seems to be a trade-off between cluttering fewer displays and increasing the display surface area. When asked about the effectiveness of the MAGE displays in reducing display count, the response was an endorsement, but not a resounding one. One SME found the information merge worked, but in a limited fashion. The second thought it was potentially useful, but needed to be better tuned to operational requirements. The last, and youngest, liked the ability to bring information together from disparate sources in one display.

SMEs seemed to like the enhanced chat function. They thought that it worked well, enhanced the ability to use the information within the different chat rooms, and liked the tagging of information and the ability to search by tags. Automated mission planning was considered useful, but in a limited fashion. There was little interest in it for normal mission execution; there simply wasn't much known beyond the next destination. The value was perceived in the timely and continuous replanning of the emergency mission.

Crews endorsed integration of MAGE into the GCS only if it were tailored and tuned to operational needs. In its current form they thought it would be useful, but not a "must have." However, they felt that careful tuning could turn the MAGE software into a valuable tool for Predator crews. They also thought MAGE function could benefit other levels of the "kill chain" above the GCS. One suggested the ability to share "screens" so the higher authority could more easily share information and make decisions more collaborative. It was pointed out that the Air Combat Command needs to see MAGE and for them to decide where such information is needed. MAGE was seen as a means of propagating a common view of the battlefield.

MAGE Stand-Alone Simulation Evaluation

An experimental, performance-based evaluation of the MAGE system was impossible without a baseline with which to compare. Self-baselined studies, using the same system with features turned on and off, are difficult to interpret. This is especially true for MAGE, in which monitoring, navigation, and mission execution are highly integrated. Thus, the approach taken was to do a subjective usability evaluation of MAGE using experienced Predator SMEs. Using these experienced operators, we were able to shed light on how well MAGE operated, which features were effective, and where future efforts would be best spent.

The study was conducted in a facility of USI, who employs and provided the Predator operator/subjects. The conducted survey asked the participants to critically evaluate each of the MAGE component technologies separately and each question had space for written comments. This took approximately a half hour to complete. Each subject took approximately 2.5 hours for training, simulation, and survey. There were four subjects on each of two days for a total of eight. All the participants were experienced pilot or sensor operators in the Predator system and are now civilians. Most of the participants had military Predator experience, though one was a civilian flight test engineer and no military experience. Questions alternated in their scale orientation, worded such that positive may tend toward both the positive and negative poles of the Likert question form. All questions were reoriented for analysis such that 1= most unfavorable, 4=neutral, and 7=most favorable; higher response means reflect more favorable disposition.

Tactical Display: Every UAV GCS we have seen at site visits, trade shows, or professional meetings has some form of spatially-oriented Tactical Display. The MAGE system is no exception and employs an augmented version of the Georgia Tech Research Institute FalconView mission planner as the basis for its Tactical Display (Figure 3). The FalconView map database serves as the backdrop for display of the UAV position, the position of other air traffic, friendly and enemy ground unit locations, weather, and restricted air spaces. The Tactical Display also displays UAV imaging tasks, weapons targets, and hosts the mission planning displays of current and proposed route alternatives.

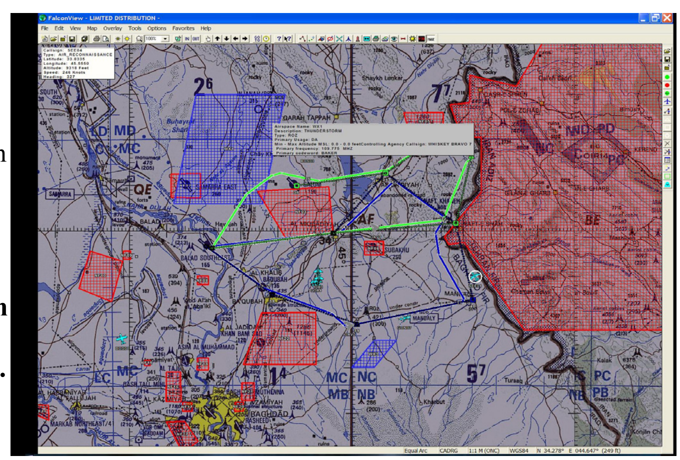


Figure 3: Tactical Display used to display UAV position, planned routes, weather, and airspaces in a spatial-map context. Information augmentation is seen in rollover and hooked textual information.

Operators were generally favorable impression of the MAGE Tactical Display. The overall assessment of the display was very effective (M_{Q11} = 5.9, SD_{Q11} = 1.6). Comments asked for a larger display to deal with clutter, better contrast control to separate symbology from the map underlay, icon management to control clutter, and a "restore" function to return to earlier configurations.

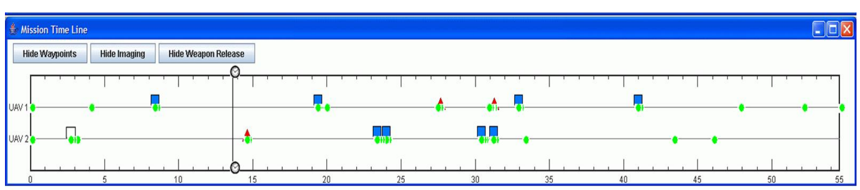
Next a series of features and representations were specifically called out for evaluation. Rollovers, which are details triggered by mouse cursor proximity, was the highest rated feature ($M_{rollovers}$ = 6.8, $SD_{rollovers}$ = 0.5). Data presentation as represented by flags, constantly visible tags that follow an icon, and hooks (where the flag is displayed to one side and the icon is marked with a halo). These were still seen as favorable (M_{flag} = 5.6, SD_{flag} =1.9; M_{hook} =, SD_{hook}), it is somewhat less enthusiastic than for the rollovers. The features assessed included airspaces, weather, blue force tracking, and other air traffic. Respectively, their means were $M_{airspace}$= 5.9, $M_{weather}$ = 5.3, $M_{blueforce}$ = 6.1, and $M_{air\ traffic}$ = 6.1 with SDs ranging from 0.6-2.1. Again, we see the dislike for automated mission planning manifest itself. One suggestion was the ability to display registered imagery from weather satellites under the symbology. Generally, the SMEs liked the rollover option and were favorable to hooks if they provided more flexibility in text placement. A preference was expressed for symbology that at least reflected types of aircraft.

Voice Recognition System: Performance of the voice recognition system was somewhat lower than had been achieved with earlier uses. The reason may have been because of the headset hardware, or because of the relative inexperience of the test users. Overall the participants were favorable to the voice recognition system. The system was graded as nearly moderately accurate (M_{Q26} = 5.8, SD_{Q26} = 0.5), better than moderately effective in context tracking (M_{Q27} = 6.3, SD_{Q27} = 0.7), and screen control was nearly moderately effective (M_{Q28} = 5.9, SD_{Q28} = 0.9). Most of the operators were familiar with FalconView's manual screen controls and preferred to use them in lieu of the voice control.

Earlier, ***SYTRONICS*** used one of the SMEs to define the MAGE fixed vocabulary. The lowest rating in voice control was in vocabulary quality, just barely appropriate (M_{Q29} = 5.3, SD_{Q29} = 2.0). This clearly indicates that with fixed vocabulary systems, enough alternative phrasing must be defined to capture more users' normal spoken language. The voice system is capable of alternative definitions, but this is time consuming and the research nature of the system did not justify the time investment. An operational system would require wider vocabulary and range of phraseology. Whatever the SME's concern about the vocabulary, they uniformly were enthusiastic about employing voice technology in the GCS (M_{Q30} = 6.3, SD_{Q30} = 0.9) and most strongly felt it could aid the Predator system (M_{Q31} = 6.8, SD_{Q31} = 0.5). This is probably in response to the manually intensive function in the Predator GCS, with operators looking to technology to reduce the manual typing workload.

Timeline Display: Timeline Displays are uncommon in operational GCSs, but a popular topic in UAV advanced development or research stations Cummings & Mitchell (2007) (Figure 4). The MAGE system is a simple temporal representation of planned mission events displayed on a linear timeline. The Timeline Display allows operators to detect co-temporal demands from the multiple vehicles. When mission replanning is requested due to task management, the alternative timelines are displayed beneath the current mission timelines so they may be compared with the current mission and other vehicles before approval of the alternative plans.

Figure 4. Timeline Display for Two UAVs with Waypoints (Green Dots), Reconnaissance Tasks (Blue Boxes), and Weapons Releases (Red Triangles)



A key issue to understanding the SME reaction to the Timeline Display is that the display assumes execution of a complete preplanned mission. This is an alien concept to Predator/Reaper crew persons, who usually operate with just a heading to maintain or do not know where or when the next object of surveillance will be directed. The Timeline Display's depiction of mission events was the lowest of any of the technologies assessed (M_{Q21} = 5.4, SD_{Q21} = 1.5). Timeline's greatest value was seen for planning an entire mission (M_{Q24} = 6.0, SD_{Q24} = 1.7). All but one SME felt it moderately or most effective; while a lone dissenter found it moderately useless. Moderately favorable averages were observed for the mission alternative comparison (M_{Q22} = 5.5, SD_{Q22} = 1.1), and value to the Predator mission (M_{Q25} = 5.5, SD_{Q25} = 1.3). The least favorable response concerned the scaling and scrolling controls of the Timeline (M_{Q23} = 4.8, SD_{Q23} = 1.9).

Chat System: The MAGE research team observed Predator crews and learned of their heavy reliance on IRC for communicating with mission essential elements around the world. However, chat continues to be used because it does not require immediate attention, it leaves a written record of communication, and reduces the chance of misunderstanding. The ***Extended Instant Messaging (xIM)*** client was developed for MAGE that integrates automatic extraction of information from the chat content, entering information into a database for later search, extraction and plotting of coordinates in a map, and highlighting high interest participants Collier, Hudson & Marshak (2007). Queries can be made through the voice recognition interface, and information can be moved from chat to other functions via an implementation of a "clipboard " shown in Figure 5.

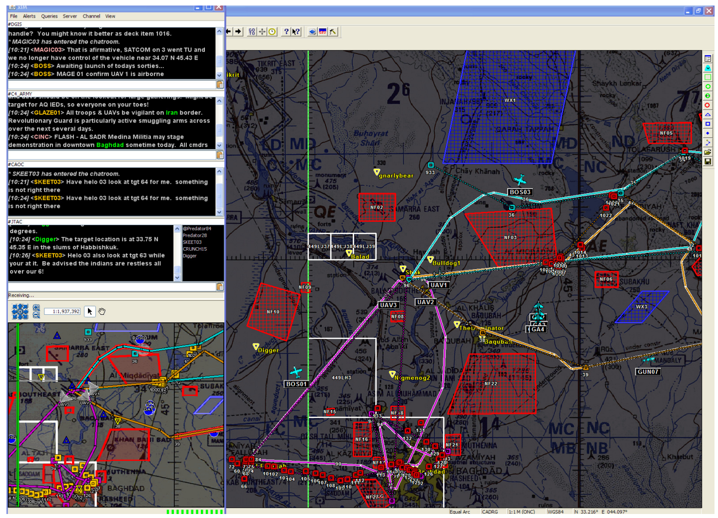


Figure 5. Extended Instant Messaging (xIM) at Left of Tactical Display with Four Chat Room Windows and Geospatial Content Display (lower left)

Again, our subject matter experts (SMEs) were largely favorable toward the chat enhanced MAGE. They found the features overall moderately useful (M_{Q32} = 6.3, SD_{Q32} = 1.0). The geo-location features of the chat system were not exploited by the test scenario, which is probable cause for the low effectiveness rating of that feature (M_{Q33} = 4.9, SD_{Q33} = 1.6). As it turns out, there was no mission demands to correlate chat content with the map, so the SMEs found the feature somewhat superfluous to the scenario. There was strong utility in highlighting of chat room call signs (M_{Q34} = 6.4, SD_{Q34} = 0.7) and making the content available to other system components via the clipboard (M_{Q35} = 6.5, SD_{Q35} = 1.1). The SMEs saw moderate utility for enhanced chat to UAS crews (M_{Q36} = 5.9, SD_{Q36} = 1.2). SME comments on xIM were entirely positive. It was judged to be a boon for situation awareness, seen as a way to recover history information, and was judged to be a time saver.

Decision Support Interface: The DSI employed Cognitive Engineering in its design and had many advanced technology components. Its design was based on a theory from cognitive psychology. Klein (1989) proposed that experts employ Recognition Primed Decision Making when faced with a decision. Experts look at the relevant information and based on their experience, "recognize" the best decision or course-of-action. The ***Decision Support System (DSS)*** uses the MAGE intelligent agent architecture to collect relevant information from Web centric sources and assemble them in a windowed workspace. One example of a DSI template is shown in Figure 6.

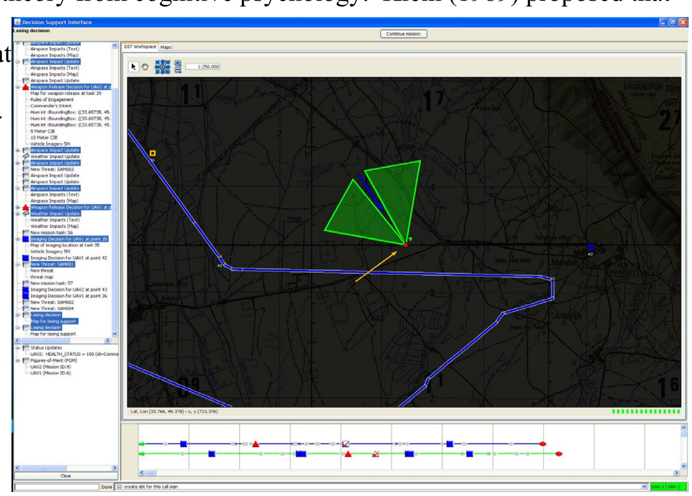


Figure 6. An Advanced DSI Display Depicting a Weapons Release Engagement with Laser Funnels (Green), Sun Direction (Yellow Arrow), Target (Red Triangle), Safe Weapons Dump Point (Green Triangle), and Approach Heading (Blue Arrow)

Information overflow and other pending decisions are displayed textually in a tree structure to the workspace's left. Additionally, users can query the system for additional information not presented by the automation. At the top of the display, the instigating event requiring intervention and the decision alternatives are presented. The decision is left to the human, but accelerated by the collection, representation, and decision expression process.

One question probed the usefulness of event triggered decision support. This fundamental principle of the DSI is that software agents would sense the mission conditions requiring operator intervention and trigger the assembly and presentation of a template. Our SMEs found this moderately useful (M_{Q37} = 6.4, SD_{Q37} = 0.9). Next, we began to evaluate the format design features individually. Identifying the instigation or trigger event at the top of the format was judged moderately useful as well (M_{Q38} = 6.1, SD_{Q38} = 1.1), eliminating any doubt about the antecedent conditions. On the left side of the format is the presentation of the question cue and data sources in tree format. Our SMEs found this only barely to moderately useful (M_{Q39} = 5.4, SD_{Q39} = 1.7). The workspace area faired slightly better, closer to moderately useful (M_{Q40} = 5.8, SD_{Q40} = 0.9). The DSI information window manager was judged to be moderately effective (M_{Q41} = 5.9, SD_{Q41} = 0.3). The SMEs did judge the query system to be moderately effective. This feature allowed operators to further information from other sources (M_{Q42} = 6.0, SD_{Q42} =0.0). Graphical representations like map mash-ups were moderately useful (M_{Q43} = 5.7, SD_{Q43} = 2.1) while text representations were rated slightly higher (M_{Q44} = 6.3, SD_{Q44} = 0.8).

Another of the new technologies evaluated was the "Sprocket" display format. This format supported choosing between mission alternatives based on a Visual Thinking design (McKim, 1972). Due to an error in the morning data collection, the format was not presented to the first four subjects and only the last three or four subjects responded in the afternoon. The three responding subjects were markedly different; two thought highly of the format rating it "highly effective" and one rated it "moderately ineffective." Individual differences are large with regard to the Sprocket format (M_{Q45} = 5.3, SD_{Q45} = 2.9). SMEs were asked to evaluate the effectiveness of the decision support format and they graded it "barely-to-moderately" effective (M_{Q46} = 5.4, SD_{Q46} = 2.0). Everyone but one SME gave it a 6-7; with one SME giving it a 1 (not at all useful). However, when asked whether the functionality could aid the Predator GCS they graded it closer to moderately useful (M_{Q47} = 5.9, SD_{Q47} = 1.0).

The response to the DSI was largely positive, though there were some suggestions to improve the implementation. Screen manipulation could be made easier by interacting with the whole data window instead of the edges as with the current implementation. Imagery should be put in track-up orientation and be centered on the target. SMEs liked the ability to mix text, graphics, and imagery in the DSS workspace. They praised the concept, indicating it would facilitate management of multiple UAVs that it brought important and high relevance information to the forefront and with the Timeline Display allowed precise management of workload.

Conclusions

Although these findings are favorable to the MAGE approach to net centric information integration, they highlight how user experience shapes acceptance of new technology. MAGE was built on the assumption of mostly autonomous UAVs, which is in contrast to the mostly manually flown Predator system. Older SME's seem biased against increasing automation. Web mash-ups like Google Maps seem to be paving the way for greater information integration in displays. Perhaps there will be greater acceptance of cognitive systems engineering of information requirements and integrated display integration.

References

Collier, J. Hudson, J. and Marshak, W. (2007). Integrating instant messaging as a net-centric information source. AUVSI's Unmanned Systems North America, August 6-9, Washington, DC.

Cummings, M.L., & Mitchell, P.J. (2007). Operator Scheduling Strategies in Supervisory Control of Multiple UAVs. *Aerospace Science and Technology,* Vol. 11, p. 339–348.

Klein, G.A., (1989) Recognition-Primed Decisions in *Advances in Man-Machine Systems Research*, W.B. Rouse, Editor., Greenwich CT: JAI Press.

McKim, R.H. (1972). Experiences in Visual Thinking. Monterey CA: Brooks/Cole.